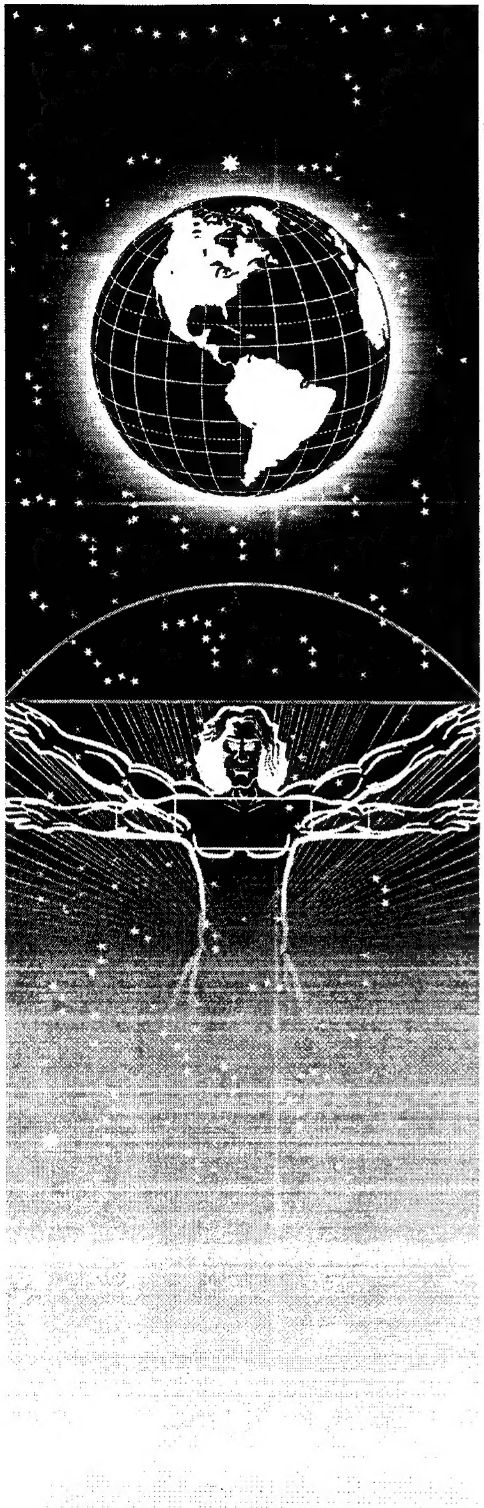
AL/HR-TP-1996-0039



UNITED STATES AIR FORCE
ARMSTRONG LABORATORY

Estimating the Training Effectiveness of Interactive Air Combat Simulation

Wayne L. Waag

Herbert H. Bell

February 1997

Approved for public release; distribution is unlimited.

Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Bldg 558
Mesa Arizona 85206-0904

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

HERBERT H. BELL
Project Scientist

ELIZABETH L. MARTIN
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

Please notify AL/HRPP, 7909 Lindbergh Drive, Brooks AFB, TX 78235-5352, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3877 or commercial (210) 536-3877.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 1997	Final Report - April 1995 to April 1996

4. TITLE AND SUBTITLE

Estimating the Training Effectiveness of Interactive Air Combat Simulation

5. FUNDING NUMBERS

PE - 62202F
PR - 1123
TA - B3
WU - 02

6. AUTHOR(S)

Wayne L. Waag
Herbert H. Bell

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Building 558
Mesa, AZ 85206-0904

8. PERFORMING ORGANIZATION REPORT NUMBER

AL/HR-TP-1996-0039

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
7909 Lindbergh Drive
Brooks Air Force Base, Texas 78235-5352

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr Herbert H. Bell, (602) 988-6561.
This paper was presented at the Advisory Group for Aerospace Research & Development (AGARD) Conference on Flight Simulation--Where are the Challenges? which was held in Braunschweig, Germany from 22-25 May 1995.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This paper documents Armstrong Laboratory's Aircrew Training Research Division (AL/HRA) participation on the Flight Vehicle Integration Panel Symposium of the AGARD Flight Conference held in May 1995. The paper is concerned with the general problem of measuring the value of simulation for combat mission training. There are a number of engineering efforts currently attempting to develop multi-player, virtual simulations that will allow soldiers, sailors, and pilots to interact with one another in a synthetic battlefield for combat mission training. This paper briefly discusses the continuation training environment that simulation must effectively complement and the various approaches for obtaining training effectiveness data for estimating the training payoff of these efforts. It summarizes the results of recent efforts conducted by AL/HRA to assess the value of combat mission simulation for continuation training of pilots. Although the results of these studies indicate high user acceptance for simulation and improved performance during the course of simulator-based training, transfer of training data has yet to be obtained. This paper was included in the AGARD Conference Proceedings 577, AGARD-CP-577, pp 37-1 to 37-8.

14. SUBJECT TERMS

Air combat; Air combat simulation; Combat mission training; Flight simulators; Interactive air combat simulation; Pilot training; Simulator-based training; Simulators; Training; Training effectiveness; Virtual simulations

15. NUMBER OF PAGES 13

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

This paper documents a paper which was presented to the Vehicle Integration Panel Symposium of the Advisory Group for Aerospace Research & Development (AGARD) Conference on Flight Simulation--Where are the challenges? which was held in Braunschweig, Germany from 22-25 May 1995. It discusses problems concerned with measuring the value of simulation for combat mission training. The paper was included in the AGARD Conference Proceedings 577, AGARD-CP-577, pp 37-1 to 37-8.

This effort was conducted under Work Unit 1123-B3-02, Tools for Assessing Situational Awareness. The principal investigator was Dr Wayne L. Waag, who recently retired. The current principal investigator is Dr Herbert H. Bell.

Estimating The Training Effectiveness of Interactive Air Combat Simulation

Wayne L. Waag, Ph.D. and Herbert H. Bell, Ph.D.

Aircrew Training Research Division
Armstrong Laboratory
6001 S. Power Rd
Mesa, AZ 85206-0904
USA

SUMMARY

This paper is concerned with the general problem of measuring the value of simulation for combat mission training. There are a number of engineering efforts currently attempting to develop multi-player, virtual simulations that will allow soldiers, sailors, and pilots to interact with one another in a synthetic battlefield for combat mission training. This paper will briefly discuss the continuation training environment that simulation must effectively complement and the various approaches for obtaining training effectiveness data for estimating the training payoff of these efforts. It will then summarize the results of recent efforts conducted by the Armstrong Laboratory to assess the value of combat mission simulation for continuation training of pilots. Although the results of these studies indicate high user acceptance for simulation and improved performance during the course of simulator-based training, transfer of training data has yet to be obtained.

1. INTRODUCTION

The United States Air Force (USAF) spends a great deal of money to develop and maintain the combat proficiency of its pilots. Most of this combat-oriented training is conducted at the operational unit as part of it's continuation training program. The basic instructional media for continuation training are the aircraft, the environment in which it operates, and the post-mission debrief. Together they provide an on-the-job training environment built around the opportunities for in-flight training. In-flight training opportunities, however, are limited by many factors (1). These factors include: peacetime training rules, resource limitations, technical constraints, and security restrictions. Each of these factors places restrictions or imposes unnatural constraints on training. Peacetime training rules impose altitude and weather restrictions, limit use of communications jamming, permit limited weapons firings, and require a minimum separation between aircraft. Resource limitations restrict the number of aircraft available for training, the number of flying hours available, and the size of the training ranges. Technical constraints limit the use of electronic warfare systems, prevent practice against an integrated air defense system, and limit the measurement of combat performance. Security restrictions prevent full employment of classified systems, communications, and tactics. These factors combine to limit the opportunities for training combat tasks at both individual and team levels.

In developing its multi-ship simulation program, the Armstrong Laboratory's Aircrew Training Research Division, in cooperation with the Air Combat Command, surveyed over 300 mission ready (MR) pilots and air weapons controllers (AWCs) to identify continuation training needs (2,3). Responses to these surveys were surprisingly similar no matter the respondent's experience level, unit, or weapon system. The consensus is that it is difficult to train the pilot and AWC to make full use of the weapon system as part of a combat team. Table 1 shows the combat training areas most frequently mentioned as needing improvement.

Table 1. Mission Activities Most Frequently Mentioned As Requiring Additional Training

Multibogey, four or more
All-aspect defense
Reaction to surface-to-air missiles
Dissimilar air combat tactics
Four-ship tactics
Reaction to air interceptors
Employment of electronic countermeasures
Chaff/flares employment

These mission areas involve the very tasks for which in-flight training is most likely to be constrained by the factors mentioned above. If anything, the negative impacts of these factors on training will increase in the future. Therefore, we must develop other training approaches that will maintain the readiness of our combat air forces. Simulation is one such approach (4). In particular, distributed interactive simulation seems especially promising since it offers the potential interactivity that characterizes the combat environment.

Because of the high cost of flight simulators and the potential consequences of inadequate training, one would assume there is an extensive research base establishing the value of training combat tasks in simulators. It is not unreasonable to ask questions such as: Was the simulator training effective? Can it be improved? How frequently is it needed? Is simulation

worth the costs? All of these questions reflect the need to evaluate the potential benefits of distributed simulation for combat-oriented training.

2. MEASURING TRAINING BENEFITS

An immediate question becomes exactly how to evaluate the benefits of simulation-based training as a means of improving combat mission performance. Bell and Waag (5) have proposed a five stage sequential model which is briefly summarized below.

Stage 1. Utility Evaluation. The objectives of the initial stage are to (a) evaluate the accuracy or fidelity of the simulation environment; and (b) to gather opinions from users concerning the potential value of the simulation for specific training applications.

Stage 2. Performance Improvement. The objective of the second stage of the evaluation is to determine the extent to which performance improves during the course of training within the simulation environment. The major challenge during this stage of the evaluation is to ensure that there is a proper means of establishing that performance has indeed improved as a result of the training. This requires the development of mission scenarios that are flown before and after the training that are similar to but not identical to missions flown during training. It also requires the development and use of measures whereby improvements in performance can be meaningfully reflected.

Stage 3. Transfer to Alternative Simulation Environment. The question of generalizability now is raised--does training transfer to another environment? While the acid test is usually considered to be transfer to the air, it is our view that a more logical intermediate step involves demonstrating transfer to other simulation environments. Recall that one of the primary justifications for multi-player air combat simulation is the ability to practice certain events under conditions that are generally not available in peacetime training environments. Because of safety restrictions, security considerations, rules of engagement, etc., peacetime exercises will always be limited in terms of their situational fidelity. For this reason, it is essential that transfer be demonstrated to another simulation environment in which a wartime environment can be created.

Stage 4. Transfer to Flight Environment. If positive transfer to a simulated wartime environment has been shown, the next stage is to show transfer to the air. Unfortunately, such a transfer test is limited by the large number of peacetime restrictions that characterize current flight operations. For this reason, a smaller sample of combat tasks would most likely have to be selected for evaluation. To whatever extent possible, the transfer test should represent a highly controlled flight environment wherein performance data can be gathered easily.

Stage 5. Extrapolation to Combat Environment. The last stage of the evaluation process attempts to answer the question of the military value of training. As might be expected, an empirical approach is not amenable for this question. Rather, a modeling approach is recommended as a means of extrapolating from simulator-based training to a combat environment. An example of such an approach is provided by Deitchman (6) in an attempt to project the impact of training into a central European type of wartime scenario. In that case, arbitrary estimates were used to represent the potential impacts of training. However, data from a systematic evaluation program, which recorded performance as a function of training, could easily be substituted into constructive models at the engagement level and the results fed into the higher level mission and campaign models. For example, training effectiveness data might show that survival is increased by an average of 25% as a result of simulator-based training. Using constructive simulations, the relative impact of such changes could be assessed in operational terms.

3. F-15 ADVANCED AIR COMBAT SIMULATION

In concert with this model, the Armstrong Laboratory has been gathering data over the past few years attempting to establish the value of simulation for air combat training. In 1988, a program was initiated with the Tactical Air Command (now Air Combat Command) to evaluate multiship air combat training using commercially available contractor facilities. In all, two utility evaluations and one simulator performance improvement study were conducted as part of this project.

These efforts used the McDonnell Aircraft Simulation facility in St. Louis, Missouri. This simulation system was designed to support engineering development. Its design and equipment typify the full mission simulator facilities developed by aircraft manufacturers in the late 1980's. Figure 1 shows the principle components of this system. Each F-15C cockpit was located in a forty-foot diameter dome which provided the pilot with a nearly full field of view. Each simulator had high fidelity aerodynamic, engine, avionics, communication, sensor, and weapon simulations. Other components included additional aircraft (either digital or manned), digitally controlled surface-to-air-threats, exercise control, debrief, and data record. A more detailed description of the basic simulation system is available (7).

Utility Evaluations. Two utility evaluations were conducted. In the first evaluation (8), 42 mission-ready F-15 pilots and 16 AWCs received four days of training. The training unit was the team comprised of two pilots (lead/wingman) plus the AWC. This team flew a variety of combat missions against an opposing force comprised of four to eight adversaries plus the adversary AWC.

Upon completion of training, pilots rated the value of both their "unit training" and the "simulation training" for 41 air-to-air tasks. The pilots felt that simulator training was much better than their current unit training for many air combat tasks

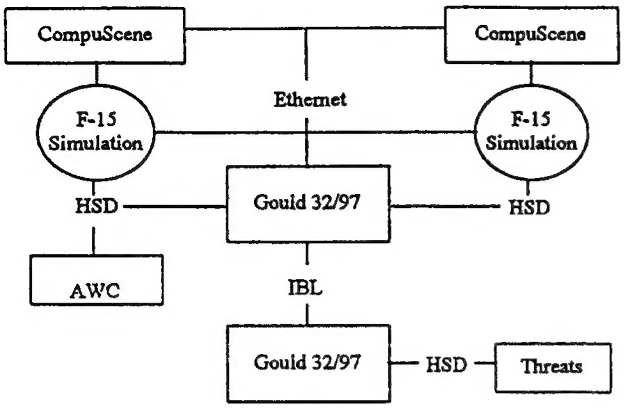


Figure 1. McDonnell Douglas Simulation Facility

including multibogey, chaff and flares employment, all-aspect defense, use of electronic countermeasures and counter-countermeasures, communications jamming, and work with the AWC. These tasks were also rated high in "need for additional training" prior to the start of simulator training. On the other hand, tasks such as air combat maneuvering (ACM), visual lookout, gun employment, and basic fighter maneuvering (BFM) were rated as better trained in their in-flight continuation training program than in the simulation. Air weapons controllers, however, rated all tasks as better trained in the simulation environment. Open-ended opinion data were also gathered, the results being quite positive toward the training.

A second evaluation, was conducted using the same procedure but with a larger sample of pilots and AWCs (3). This evaluation produced essentially the same results. Based on the high user acceptance demonstrated during these utility evaluations, Air Combat Command continued this program under its own sponsorship.

Performance Improvement. In the third study, again using the same facility, in-simulator learning was also shown, in addition to positive user opinion. Subjects consisted of 16 teams, each team being made up of two pilots and an AWC. Each of the elements flew controlled offensive and defensive scenarios "before" and "after" three days of intensive simulation training. Digital data as well as videotapes of displays used for replay and debriefing purposes were archived for later analysis.

Preliminary analyses reveals that post-training mission performance is significantly higher than pretraining performance. Figure 2 shows the mean value of several pre- and post-training mission performance indicators for defensive counterair missions. The data clearly indicate that the probability of mission success (i.e., no enemy strikers to target) and F-15 survival increased during the course of the simulator training ($p < .05$). In addition, weighted exchange ratios, reflecting the efficiency of mission accomplishment, also increased as a function of training ($p < .05$).

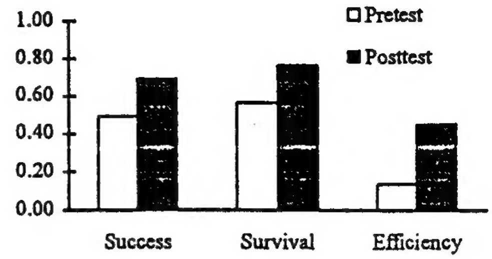


Figure 2. Comparison of Pre/Post Test Performance

4. F-15 SITUATION AWARENESS STUDY

In 1991, the US Air Force Chief of Staff posed a series of questions concerning situation awareness (SA). First of all, what is SA? Can it be objectively measured? Is SA learned or does it represent a basic ability or characteristic that some pilots have and others do not? From a research standpoint, these questions translate into issues of measurement, selection, and training. The Armstrong Laboratory was subsequently tasked with providing research answers to these questions. A research investigation was initiated that had three goals: first, to develop and validate tools for reliably measuring SA; second, to identify basic cognitive and psychomotor abilities that are associated with pilots judged to have good SA; and third, to determine if SA can be learned, and if so, to identify areas where cost-effective training tools might be developed and employed. An overview of the investigation can be found in McMillan, Bushman, and Judge (9).

The general approach was to first develop criterion measures of SA based upon performance ratings collected within an operational flying environment. The results of this part of the study can be found in Waag and Houck (10). These measures were necessary for two reasons. First, they would serve as criterion measures against which to validate a battery of basic ability tests considered relevant to SA, thereby addressing the question of basic human abilities. The results of this part of the study can be found in Carretta and Ree (11). Second, these measures would serve as a means of selecting a sample of pilots who would participate in a simulation phase of the effort. During that phase, simulated air combat mission scenarios were developed for assessing SA and objective measures of performance gathered in an attempt to determine those characteristics that distinguish pilots with good SA. These data would be used to identify areas where training tools might be developed. We now summarize the results of the third phase of the program, namely the use of simulation as a tool for measuring and training SA. The complete findings are presented in Waag, Houck, Greschke and Raspotnik (12).

Method. A total of 40 MR F-15 pilots, who were flight lead-qualified served as subjects. An additional 23 MR F-15 pilots served as wingmen throughout the data collection. The simulated combat missions were flown using the Armstrong

Laboratory's multiship simulation facility (MULTIRAD) located at Williams AFB (WAFB), AZ. The major components of the simulation system are shown in Figure 3. These components represent independent subsystems operating as part of a secure distributed simulation network. This local area network was connected to the air weapons controller simulator (AESOP) at Brooks Air Force Base (BAFB), TX by a dedicated T-1 telephone line. Additional details concerning the basic simulation architecture and components are available (13,14,15).

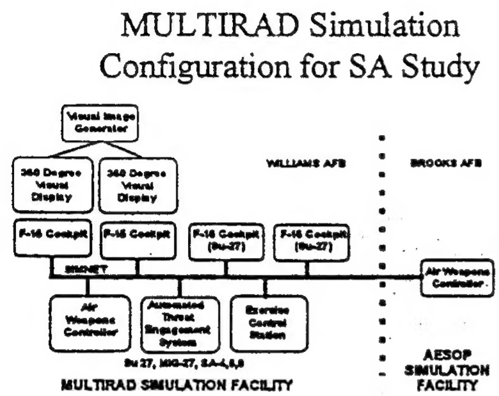


Figure 3. Armstrong Laboratory Simulation Facility

The manned flight simulators consisted of two F-15C simulators and two F-16 simulators. The F-15C simulators had high fidelity aerodynamic, engine, avionics, radio, sensor, and weapons simulations. Each F-15C simulator was equipped with an out-the-window visual display system covering approximately 360 deg horizontal by 200 deg vertical. The external visual scene was created using computer-generated imagery. The lower fidelity, manned F-16 simulators played the role of enemy aircraft in conjunction with computer-controlled adversaries. The visual and electronic signatures of these F-16 simulators were modified so that they appeared as the appropriate threat aircraft. Each F-16 simulator was equipped with a single channel of out-the-window visual imagery covering approximately 45 deg horizontal by 45 deg vertical. A manned AWC provided the F-15C pilots with appropriate threat information and warnings. Depending upon the availability of qualified AWCs and equipment status, the AWC was either located at WAFB or BAFB. In either case, the AWC had a realistic simulation of the appropriate AWC console and communicated with the F-15C pilots by radio.

The primary approach taken toward the measurement of SA was through scenario manipulation and observation of subsequent performance. A week-long SA "evaluation" exercise was constructed that consisted of 9 sorties with 4 engagements per sortie. Sorties were arranged in a building block manner. Over the week, engagements increased in complexity in terms of numbers of adversaries, enemy tactics, lethality of ground threats, AWC support, etc.

The same two subject-matter experts (SMEs) were used throughout the year-long data collection effort. Upon completion of the mission, they discussed each engagement, and completed a consensus performance rating scale consisting of the 24 behavioral indicators of SA related to F-15 mission performance. A variety of other data were also gathered and archived, including mission events and outcomes, digital data passed over the network, videos used for debriefing, eye movement data recorded on the last mission, and finally, "critiques" of the simulation and opinions regarding its potential for training. Two types of user opinion data were gathered. First, pilots rated the training benefit for various pilot experience levels. And second, pilots completed an open-ended questionnaire pertaining to the overall value of the simulation and how it might best be used.

Findings. The results of the ratings of potential training benefits are provided in Figure 3. These data clearly indicate that positive opinions were expressed by the study participants on the value of this type of simulation for training. The potential training was considered beneficial for all levels of qualification. It is of interest to note that training was considered highly beneficial for four-ship flight leads, despite the fact that the MULTIRAD simulation facility provided training for only a flight lead and wingman. As expected, higher benefit ratings were given to pilots upgrading into a given qualification level.

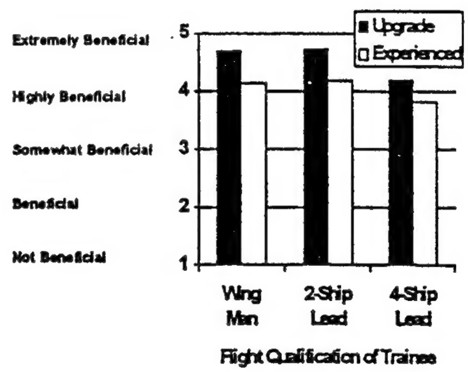


Figure 4. Rated Benefits of Simulation Training

Opinions expressed in the open-ended questionnaire were also quite positive. Although qualitative, they provide additional insight into the potential focus of training using multiship simulation and how it might be employed. In particular, mention was made of using such training as a means of enhancing both situation assessment and decision-making skills. It was also frequently noted that there was tremendous value in learning flight leadership and resource management skills. In terms of the location of such simulation, the overwhelming consensus was that they would be of most value within the operational units. This was not too surprising since each unit now has the operational version of the cockpits used in the present investigation. However, they are stand-alone and non-visual, and as such their training capability is fairly

limited. In contrast, the networking of such devices within a realistic combat environment increases the potential greatly. The bottom line from the utility data is that the participants considered multiship simulation as a tool with high training potential.

It should be pointed out that it was never the intent, at the outset of the study, to demonstrate performance improvements. It must be emphasized that the sole purpose was to develop a set of simulation scenarios that could be used to assess SA within a combat environment. As such, normal training interventions were not permitted. For example, during the debrief, pilots were permitted to only view their own in-cockpit displays and not the planned view display. Moreover, the two SMEs were not permitted to provide any type of feedback to the pilots regarding their performance. However, data from the ninth mission did permit some comparison since identical scenarios had been flown earlier in the week. The ninth mission was designated the "eye track" mission in which eye movement data were recorded.

Two scenarios, a 2 V 2 defensive counterair (DCA) mission and a 2 V 4 offensive counterair (OCA) mission, were flown during the middle of the week and then again on the last mission. A comparison of performance is presented in Figure 5. In both cases, performance on the last mission was improved. However, only the 2 V 2 DCA mission was found to be statistically significant. When such data are coupled with the very strong pilot opinions that they had received valuable training, it seems reasonably safe to conclude that learning had occurred over the week.

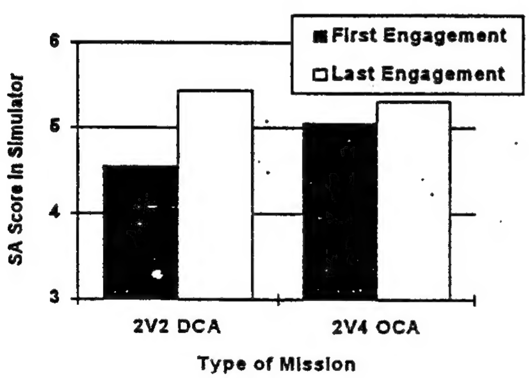


Figure 5. Effects of Practice on Observer SA Ratings

5. MULTI-SERVICE DISTRIBUTED TRAINING TESTBED

The Armstrong Laboratory is currently working with the Army Research Institute and the Naval Air Warfare Center to develop a training testbed that can be used to assess the value of Distributed Interactive Simulation for multi-service training. This effort, initially sponsored by the Defense Modeling and Simulation Office, links service-developed training systems together to create a Multi-Service Distributed Training Testbed (MDT2). MDT2 provides a common, virtual environment that is being used to support training research involving collective tasks.

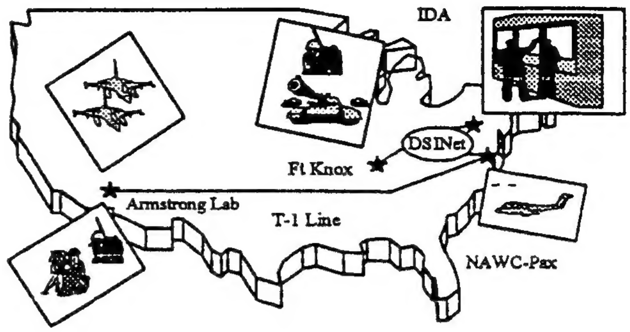


Figure 7. Multi-Service Distributed Training Testbed

Currently, MDT2 is focusing on the planning and execution of close air support (CAS) at the engagement level. The goal is to establish a virtual training environment. The participants in this virtual training environment will be soldiers, marines, and pilots executing their unique combat tasks in virtual simulators at their individual service training sites. This environment will allow collective training to occur which involves both unit and task force components.

The initial training utility evaluations of MDT2 were conducted in May of 1994 and February of 1995. For the February evaluation, four sites were interconnected using Distributed Interactive Simulation Protocol 2.0, version 3. These sites were the Institute for Defense Analysis in Alexandria, Virginia; Mounted Warfare Testbed at Ft. Knox, Kentucky; the Manned Flight Simulation Facility at Patuxent River, Maryland; and the Armstrong Laboratory at Mesa, Arizona. This simulation network is illustrated in Figure 7, while information regarding each of the sites is summarized in Table 2.

Table 2. Simulation Testbed Components

Location	Simulators	Role
Armstrong Laboratory	2 F-16	Close Air Support
	1 Laser Designator	Scout and Target Designation
Naval Air Warfare Center	OV-10	Forward Air Control
Mounted Warfare Testbed	M-1A1	Ground Maneuver
	M-2	Battalion Tactical Operations Center
Institute for Defense Analysis	None	Stealth Display Data Record

The results of this training effectiveness evaluation indicated that mission ready fighter pilots, airborne forward air controllers, and ground combatant felt that the simulation

provide significant training benefits. In addition, trained observers and subject matter experts monitored that performance of selected mission tasks throughout the training period. The data indicated that mission performance for each of the components, air and ground, increased over the four days of simulated combat. The high user acceptance and increase in performance are most likely due to unique ability that MDT2 provides for planning, execution, and review of close air support as an integral part of the ground commander's battle plan.

6. DISCUSSION

The results obtained from these efforts provide strong user belief in the value of interactive air combat simulation. From the user's perspective, the data are very clear regarding the potential value of such simulation for training. User's consistently report that such simulations are an enhancement to their current mission training. Although such subjective evidence is often considered suspect from a scientific perspective, it is nevertheless an absolute prerequisite for effective training. Unless there is user acceptance, the resulting training will be of marginal value regardless of the device's inherent potential.

In addition to the opinion data, there is evidence that performance did improve within the simulation environment. Performance improvements were demonstrated in the F-15 Advanced Air Combat Simulation, the F-15 Situation Awareness Study, and the Multi-Service Distributed Training Testbed. These data combined with the fact that the study participants expressed opinions to the effect that their proficiency had improved leave little doubt that learning had occurred.

Although the data clearly indicate that the end user expresses very positive opinions toward the value of multiship simulation and that learning occurs, there still remains the issue of transfer of training. Does such training transfer to other simulation environments (Stage 3 of the Evaluation Model) and does it transfer to the real world (Stage 4 of the Evaluation Model)? The data gathered in this study do not bear upon these issues.

The question becomes, "are transfer of training data needed?" While no one would argue the desirability of having such data, there are practical issues which seriously question the advisability of conducting such studies. Lack of experimental control, insufficient sample sizes, insufficient training time in the simulator, insufficient time for evaluating transfer in the air, insensitive measures, etc. are problems that plague the conduct of any transfer of training evaluation (16). In fact, one can argue that it is virtually impossible to conduct a well-controlled transfer test within an operational military environment.

This inability to adequately control such evaluations perhaps has its greatest impact on the interpretation of findings, particularly when these findings show no transfer effects or fairly small transfer effects. The empirically obtained outcome for any transfer of training experiment is one of three possible outcomes; positive transfer, no transfer, or negative transfer. Similarly, the true effect of training is one of the same three possibilities. The problem is to infer the true state from the obtained outcome. This inferential process works quite well when statistically significant outcomes confirm expectations. Unfortunately, the inference process does not work as easily when little or no transfer is obtained as a result of training. Now the investigator must decide between two possibilities. Indeed, the training may have little of no effect on performance. Or, the effects may be much larger, but because of methodological problems inherent in conducting transfer of training experiments, they are masked. Although we do not know the true effects of training, we generally attempt to "explain away" any lack of positive effects and attribute it to these "methodological problems", especially if there are other data such as expert opinion that suggest the training to be beneficial.

A good case in point is a study by Pohlmann & Reed (17) that failed to show positive transfer effects for air combat maneuvering (ACM) training in the Simulator for Air-to-Air Combat (SAAC). Do we believe that simulator training does not improve air-to-air performance? Probably not, since we have other evidence suggesting the training to be beneficial. This evidence includes positive end-of-course critiques indicating that such training in the SAAC was some of the best air-to-air training pilots had ever received, in-simulator performance improvements (18), and positive transfer of training in another experiment (19). The study failing to demonstrate positive transfer had one potentially serious limitation in that instructor ratings were used as a measure of performance. Such measures have been shown to be quite insensitive in other air combat domains. For example, a study by Gray & Fuller (20) which demonstrated significant transfer of training in terms of bombing accuracy, also used instructor ratings of performance in the air. Interestingly enough, the rating data showed no effects of simulator pretraining despite large differences in objective measures of weapons delivery. So it seems at least plausible that the failure to show any effect in the Pohlmann & Reed (17) study may have been due largely to the measures that were used. For this reason and the fact that we have other evidence suggesting the training to be valuable, we can make the case to simply "dismiss" these findings.

At this point, we have a paradox emerging. On the one hand, we have made the argument that the transfer of training evaluation is the only <u>sufficient</u> test for establishing training effectiveness. On the other hand, we have also shown that we tend to dismiss those studies failing to demonstrate positive transfer when we have other data, which is usually expert opinion, suggesting the training to be effective. In such instances we attribute the lack of positive transfer effects to one or more of those "methodological problems" which always exist in the conduct of such evaluations within an operational

military training environment. If we are willing to explain away our inability to demonstrate training effectiveness, the question becomes, "why conduct the transfer evaluation?"

Since there is currently no definitive data, the question of training benefits of interactive air combat simulation is largely answered by one's personal view of simulation and one's willingness to generalize from previous investigations of transfer in other domains. For the "believer," including the authors of this paper, the evidence to date is strong enough to warrant the conclusion that training will be effective. In fact, given the previous transfer of training research that has already been conducted (16,5) there is little reason to suspect that such training within a multiship simulation environment would not have a positive effect upon subsequent performance in the air. Consequently, there is no compelling reason to conduct transfer of training studies within the air combat environment. However, for the "skeptic," no definitive evidence has been presented and the question remains unanswered.

7. REFERENCES

1. United States Air Force Scientific Advisory Board. (1992, December). Report of the Support Panel of the Air Force Scientific Advisory Board Ad Hoc Committee on Concepts and Technologies to Support Global Reach - Global Power 1995 - 2020. Washington, DC: Author.

2. Gray, T.H., Edwards, B.J., & Andrews, D.H. (1993). A survey of F-16 squadron level training in PACAF (AL-TR-1993-0041). Williams Air Force Base, AZ: Armstrong Laboratory, Aircrew Training Research Division.

3. Houck, M.R., Thomas, G.S., & Bell, H.H. (1991). Training evaluation of the F-15 advanced air combat simulation (AL-TP-1991-0047). Williams Air Force Base, AZ: Armstrong Laboratory, Aircrew Training Research Division.

4. Alluisi, E.A. (1991). The development of technology for collective training: SIMNET, a case history. Human Factors, 33, 343-362.

5. Bell, H.H. & Waag, W.L. (1994). Evaluating the Effectiveness of Flight Simulators for Training Combat Skills: A Review. Submitted to International Journal of Aviation Psychology.

6. Deitchman, S.J. (1988). Preliminary exploration of the use of a warfare simulation model to examine the military value of training (IDA Paper P-2094). Alexandria, VA: Institute for Defense Analysis.

7. McDonnell, G.W., Broeder, R.F., & Cutak, R.J. (1989). Multi-ship air combat simulation. In Proceeding of the 11th Interservice/Industry Training Systems Conference. Ft. Worth, TX: National Security Industrial Association.

8. Thomas, G.S., Houck, M.R. & Bell, H.H. (1990). Training evaluation of air combat simulation (AFHRL-TR-90-30). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.

9. McMillan, G.R., Bushman, J. & Judge, C.L.A. (1995) Evaluating pilot situational awareness in an operational environment. In AGARD-CP-575, Situation Awareness: Limitations and Enhancement in the Aviation Environment, Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research and Development.

10. Waag, W.L. & Houck, M.R. (1995) Development of criterion measures of situation awareness for use in operational fighter squadrons. In AGARD-CP-575, Situation Awareness: Limitations and Enhancement in the Aviation Environment, Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research and Development.

11. Carretta, T.R. & Ree, M.J. (1995) Determinants of situational awareness in U.S. Air Force F-15 pilots. In AGARD-CP-575, Situation Awareness: Limitations and Enhancement in the Aviation Environment, Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research and Development.

12. Waag, W.L., Houck, M.R., Greschke, D.A. & Raspotnik, W.B. (1995) Use of multiship simulation as a tool for measuring and training situation awareness. In AGARD-CP-575, Situation Awareness: Limitations and Enhancement in the Aviation Environment, Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research and Development.

13. Gehl, T.L., Rogers, R.L., Miller, M.A., & Rakolta, J. (1993). Interdependence of training utility and network performance using the Armstrong Laboratory Multiship Research and Development System. In, Proceedings of 15th Industry/Interservice Training Systems Conference, Orlando, FL: National Security Industrial Association.

14. Platt, P. & Crane, P. (1993). Development, test, and evaluation of a multiship simulation system for air combat training. In, Proceedings of 15th Industry/Interservice Training Systems Conference, Orlando, FL: National Security Industrial Association.

15. Rogers, B. (1992). Tactical air threat system for a distributed simulation network. In, Proceedings of 14th Industry/Interservice Training Systems Conference, San Antonio, TX: National Security Industrial Association.

16. Waag, W.L. (1981). Training effectiveness of visual and motion simulation. (AFHRL-TR-79-72). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.

17. Pohlmann, L.D. & Reed, J.C. (1978). Air-to-air combat skills: contribution of platform to initial training (AFHRL-TR-

78-53). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.

18. Robinson, J.C., Eubanks, J.L. & Eddowes, E.E. (1981). Evaluation of pilot air combat maneuvering performance changes during TAC ACES training. Nellis AFB NV: USAF Tactical Fighter Weapons Center.

19. Jenkins, D.H. (1982). Simulation training effectiveness evaluation (TAC Project 79Y-001F). Nellis AFB, NV: Tactical Fighter Weapons Center.

20. Gray, T.H. & Fuller, R.R. (1977). Effects of simulator training and platform motion on air-to-surface weapons delivery training (AFHRL-TR-77-29). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.